



Home › Statistics › Detailed methodology information › Concepts, sources and methods  
› Labour Statistics: Concepts, Sources and Methods › 2023 › Methods: Four pillars of labour statistics  
› Administrative data › Linked Employer-Employee Dataset (LEED)

## ✓ Latest release

# Linked Employer-Employee Dataset (LEED)

## Labour Statistics: Concepts, Sources and Methods

Reference period: 2023

Released 24/11/2023

### Sections

#### ▼ [Labour Statistics: Concepts, Sources and Methods, 2023](#)

[Overview](#)

▶ [Concepts and sources](#)

▼ [Methods: Four pillars of labour statistics](#)

▶ [Household surveys](#)

[Business surveys](#)

▼ [Administrative data](#)

[Linked Employer-Employee Dataset \(LEED\)](#)

### On this page

[LEED overview](#)

[Person file](#)

[Jobs file](#)

[Employer file](#)

[LEED outputs](#)

[Single Touch Payroll \(STP\)](#)

▶ [Australian Labour Account](#)

[Glossary](#)

[Acronyms](#)

Close

The Linked Employer Employee Dataset (LEED) is a cross-sectional database which is built using Australian Tax Office (ATO) administrative data linked to ABS [Business Longitudinal Analytical Data Environment \(BLADE\) \(/about/data-services/data-integration/integrated-data/business-longitudinal-analysis-data-environment-blade\)](#).

The LEED enables simultaneous analysis of met supply and demand in the Australian labour market, through:

- providing supplementary labour statistics and facilitates labour market research at industry and regional levels.
- enabling analysis of the Australian labour market at macro and micro levels;
- enabling analysis of how specific events impact employees and employers;
- helping to understand structural changes in the labour market.

The LEED consists of three cross-sectional files:

- a person file;
- a jobs file; and
- an employer file.

The LEED associates information about a person with information about their employing business. This is done by establishing the existence of a job. An employed person can have one or more jobs throughout the year with one or more employers, some of which may be held concurrently with others. A job can be created either by an employing business or the personal enterprise of the individual (an owner manager).

## **LEED overview**

### **Scope**

The LEED contains information for all persons who interacted with the Australian taxation system since the 2011-12 financial year. The LEED covers all persons who either:

- submitted an individual tax return (ITR); or
- had an Income Statement (previously Pay As you Go (PAYG) payment summary) issued by an employer and then remitted to the ATO.

Employees who did not submit a tax return and have not provided their Tax File Number to their employer will not appear in the LEED. Owner managers of unincorporated enterprises (OMUEs) who did not submit an ITR are also excluded.

The LEED includes all sources of income, regardless of whether the income provider is

based within Australia's economic territory.

## **Migrant data**

From 2022, migrant data were added to the LEED. The migrant data used in LEED are sourced from the Person Level Integrated Dataset (PLIDA), formerly known as the Multi-Agency Data Integration Project (MADIP).

The migrant data are a suite of administrative datasets (visa grants and settlements database) from the Department of Home Affairs. These data pertain to permanent migrants and temporary entrants to Australia.

## **Integration methodology**

LEED links jobs to employers and employed persons are linked to employers via the jobs they hold.

Initial data cleaning is undertaken to remove duplicate and erroneous records. In particular, job records are repaired to minimise the impact of administrative noise on output statistics, such as annual income statements issued in two separate parts.

Before the linkage takes place, an input job level file is created largely based on the income statement file. This file is also enhanced with job records derived using ITR information, to cover jobs without income statement information, such as OMUE jobs. Data quality of this file is also enhanced using occupation information from ITR, and the best available age, sex, and geographic information between the Income Statement, ITR and Client Register (CR) data.

Jobs are integrated with the employer by one of two methods. The method is dependent on which part of the business population on the ABS Business Register the employer is grouped into.

- Non-profiled population (businesses with a simple structure): a deterministic approach using the Australian Business Number (ABN).
- Profiled population (businesses with a complex structure): a more detailed approach to linking is used, detailed below.

## **Profiled population linking**

Where an employer is part of the profiled population, the relevant jobs are assigned to type of activity units (TAUs) based on a logistic regression model developed using Census data. The model references independent variables common to both Census and personal income tax data, including sex, age, occupation, and region of usual residence. These are used to predict the industry of employment, which conceptually aligns to a type of activity unit.

Where an employee has multiple job relationships with the same reporting ABN in an enterprise group, each job relationship is assigned to the same type of activity unit.

Based on the model, each job record is assigned a probability of being in each of the type of activity units present in the employing enterprise group. Iterative random assignment is undertaken using these probabilities until employment benchmarks are met. Benchmarks are based on Quarterly Business Indicators Survey (QBIS) data where a unit is in scope. BLADE employment levels are substituted where QBIS data is not available, otherwise no benchmarking is done.

The above process is applied to link the different input datasets for each financial year. Records have not been integrated across years and therefore, the LEED is a cross-sectional database and is not longitudinal.

### **Integrating migrants data**

Personal identifiers were used to first integrate the migrant data with the ATO's Client Register data and then integrated into LEED. This enables more detailed analysis of labour market and fiscal contributions of migrants to the economy, allowing policy makers and researchers to better understand the migrant experience and their economic contribution to Australia.

ABS data integration practices comply with the High-Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes. For further information see - [Keeping integrated data safe \(/about/data-services/data-integration/keeping-integrated-data-safe\)](/about/data-services/data-integration/keeping-integrated-data-safe).

### **Legislative environment**

The LEED incorporates:

- person level ITR data, job level income statement data and Client Register data supplied by the ATO to the ABS under the Taxation Administration Act 1953 - which requires that such data is only used for the purpose of administering the Census and Statistics Act 1905; and
- employer level data that include the ABS's BLADE data and the [ABS Business Register \(https://www.abs.gov.au/ausstats/abs@.nsf/dossbytitle/AC79D33ED6045E88CA25706E0074E77A?OpenDocument\)](https://www.abs.gov.au/ausstats/abs@.nsf/dossbytitle/AC79D33ED6045E88CA25706E0074E77A?OpenDocument) data supplied by the Registrar of Australian Business Register (ABR) to the ABS under A New Tax System (Australian Business Number) Act 1999 - which requires that such data is only used for the purpose of carrying out the functions of the ABS.

The data limitations or weakness outlined here are in the context of using the data for statistical purposes, and not related to the ability of the data to support the ATO's or ABR's

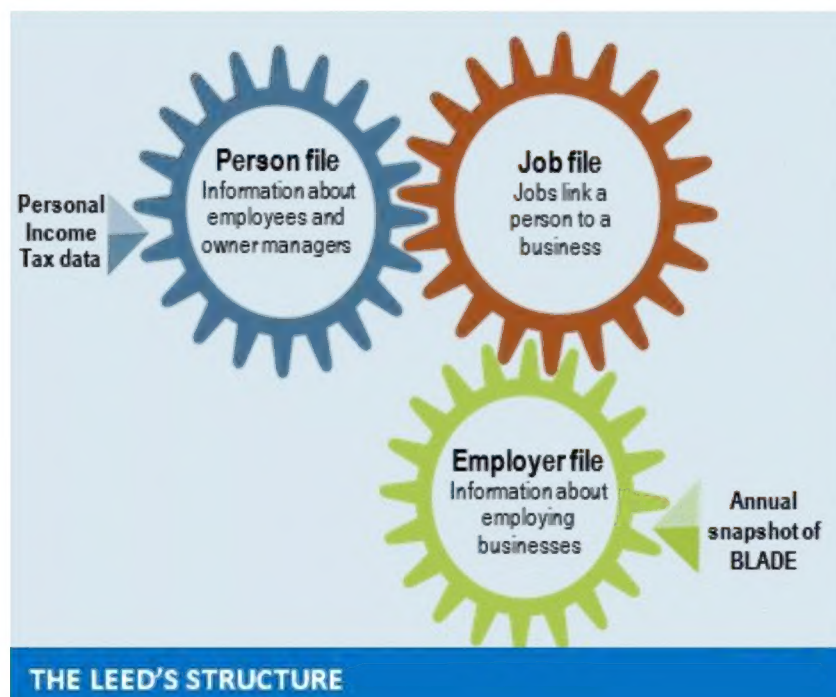


core operational requirements.

Legislative requirements to ensure privacy and secrecy of these data have been followed. In accordance with the Census and Statistics Act 1905, results have been confidentialised to ensure they are not likely to enable identification of a particular person or organisation. All personal information is handled in accordance with the Australian Privacy Principles contained in the Privacy Act 1988.

All personal income tax statistics were analysed in de-identified form with no home address or date of birth included in LEED input files. Addresses were coded to the [Australian Statistical Geography Standard \(ASGS\) \(/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/latest-release\)](https://statistics/standards/australian-statistical-geography-standard-asgs-edition-3/latest-release) and date of birth was converted to an age at 30 June of the reference year prior to data provision.

**The LEED is comprised of a person file, a job file and an employer file**



## Person file

Each person file contains data for all persons who either submitted an Individual Tax Return (ITR) or who were identifiable on an income statement in the reference year. Each record includes de-identified demographic and geographic data, and aggregate income information.

Employed persons may be either employees (including Owner Manager of Incorporated Enterprises or OMIEs), Owner Managers of Unincorporated Enterprises (OMUEs), or both.

Employees are identified by the presence of aggregate employee income and at least one linked employee job.

Employees who have not submitted an ITR but who have provided their Tax File Number to their employer are imputed from income statement data.

OMUEs are identified by the presence of any of the own unincorporated business income types and a linked OMUE job.

Tax lodgers who are not employees or owner managers (such as persons with only investment incomes) are included on the person file to support statistical analysis that requires a more complete view of the tax lodger population.

## **Jobs file**

The jobs file is a complete list of the job relationships held at any time during the reference year. It is constructed primarily from income statement data. Income statements describe the payments made to an individual by an employer within a financial year. Conceptually, income statement data should include most employee/employer job relationships. OMUE jobs are derived from ITR data and are added to the jobs file, some of these link to businesses in the Business Longitudinal Analysis Data Environment (BLADE).

In some cases a synthetic employee job record has been created based on information in the person file. This has occurred when a person has recorded wage or salary information that cannot be identified in income statement data. Sometimes, an employee job may not be able to be linked to an employing organisation due to reporting errors or missing information.

A person can hold several jobs during the year, either concurrently (as a multiple job-holder) or consecutively. For a person who is an employee of several employers, each relationship is listed as a separate job. Due to data limitations, only one self-employment job can be recorded for any OMUE even if a person owns and manage more than one enterprise. In the LEED, an OMUE can hold other jobs as an employee.

The LEED jobs file excludes voluntary jobs and unpaid contributing family worker jobs.

## **Employer file**

In the LEED, an employer is any legal entity in the non-profiled population that is linked to a job; and any type of activity unit in the profiled population that is linked to a job.

The employer file contains business units present in BLADE that could be linked to a job, as well as unincorporated entities. Some unincorporated entities are identified in personal income tax data and are not otherwise included in BLADE or cannot be identified in BLADE. Industry and several other employer variables are not available for these unincorporated entities, except from 2017-18, where industry information in their ITR has been used if available.

## **LEED outputs**

### **Key outputs**

The LEED provides cross-sectional information relating to employees and owner managers of unincorporated enterprises

Key data/series include:

- Employed persons and their jobs (employees and owner managers of unincorporated enterprises)
- Multiple job holders
- Income at job and person levels
- Regional spotlights of jobs and employed persons

Other data includes (but is not limited to):

For people with income:

- Income types: Total, Employee, Investment, Own unincorporated business, Superannuation
- Counts of earners
- Distributional information: mean, median, quartiles, percentile ratios, gini coefficient, income share
- Geography - region of residence (at State and Territory, Local Government Area, Statistical Area 4, 3, and 2 levels)
- Demographic information: age, sex
- Migrant characteristics: visa, year of arrival, applicant status

In addition, for persons with jobs:

- Counts: Employed persons, Jobs, Employees, Owner-Managers of Unincorporated Enterprises, Multiple job holders
- Status in employment: Employee, Owner-manager of Unincorporated Enterprise
- Income: Employment, Employee, Own Unincorporated Business, Duration adjusted

income per job (annualised)

- Detailed occupation and skill levels of persons
- Detailed industry of job
- Sector (public/private)
- Number of jobs held (employee jobs and owner manager of unincorporated enterprise jobs)
- Duration of jobs
- Concurrent and non-concurrent jobs

Information relating to employers:

- Employment size
- Detailed industry of business activity
- Type of legal organisation (TOLO)
- Institutional sector (SISCA)

## Statistical releases

LEED data is disseminated through the publications listed below. Additional data is available through Customised Data Requests.

[Jobs in Australia \(/statistics/labour/earnings-and-work-hours/jobs-australia/latest-release\)](/statistics/labour/earnings-and-work-hours/jobs-australia/latest-release)

Frequency: Annual, from 2011-12

Jobs in Australia (JIA) provides aggregate statistics from the Linked Employer-Employee Dataset. It provides information about filled jobs in Australia, the people who hold them, and their employers. JIA provides data across 2,288 Statistical Areas as well as Local Government Areas.

[Personal Income in Australia \(/statistics/labour/earnings-and-work-hours/personal-income-australia/latest-release\)](/statistics/labour/earnings-and-work-hours/personal-income-australia/latest-release)

Frequency: Annual, from 2011-12

Personal Income in Australia (PIIA) provides a comprehensive range of income indicators across small geographic areas.

[Tablebuilder: Jobs and Income of Employed Persons \(/statistics/microdata-tablebuilder/available-microdata-tablebuilder/jobs-and-income-employed-persons\)](/statistics/microdata-tablebuilder/available-microdata-tablebuilder/jobs-and-income-employed-persons)

Frequency: Annual, from 2011-12

Release of LEED data for employed persons through TableBuilder. This enables users to build their own customised tables from the Linked Employer-Employee Dataset microdata, including for State and Commonwealth Electoral Divisions.



